Uncertainty Quantification in Viscous Hypersonic Flows using Gradient Information and Surrogate Modeling

Brian A. Lockwood,* Markus P. Rumpfkeil,† Wataru Yamazaki,‡ and Dimitri J. Mavriplis §
University of Wyoming, Laramie, 82071, USA

In this paper, the use of gradient observations in conjunction with surrogate models for the purposes of uncertainty quantification within the context of viscous hypersonic flows is examined. The approach is presented within the context of both perfect gas and non-equilibrium real gas simulations and can be used to quantify the uncertainty in an output of interest due to the uncertainty associated with various model and design parameters. The gradient of an objective is calculated via a discrete adjoint approach. By using an adjoint based approach, the sensitivity of an objective to a large number of input parameters can be calculated in an efficient and timely manner. With these sensitivity derivatives, the uncertainty in an objective due to input parameters is calculated in various ways. Initially, first-order methods, such as the method of moments and linear extrapolation, are used to represent the design space and calculate relevant output statistics. In order to improve upon these first-order approaches, Kriging and gradient enhanced Kriging models are created for the function space and serve as a basis for additional inexpensive Monte Carlo sampling. The probability distribution functions and statistics generated by these linear and Kriging based methods compare favorably with the results of nonlinear Monte Carlo sampling and can serve as a basis for further exploration of uncertainty quantification in hypersonic flows.

I. Introduction and Motivation

The increasing reliance on numerical simulation for the design and analysis of physical systems has lead to a desire for uncertainty analysis in order to assess the quality of simulations and apply confidence bounds to outputs. In particular, uncertainty quantification is especially important for problems in which validation data is difficult or impossible to obtain, as is the case with hypersonic flows. The uncertainty in simulation outputs can arise from various sources, such as measurement errors, modeling inadequacies,[1] or manufacturing tolerances.[2] Due to these factors, predicting a specific output for a given set of input parameters can be insufficient and a distribution or interval bound of outputs based on the uncertainties inherent in the simulation parameters is required. For hypersonic flows, Monte Carlo sampling is typically used to quantify uncertainty in outputs of interest, requiring thousands of flow solves to collect the appropriate statistics.[3,4] In order to reduce the computational cost of uncertainty quantification, gradient information can be leveraged to reduce the number of samples required or provide samples at reduced cost.[5–8] Most notably, adjoint methods may be used to provide the derivatives of a single output with respect to a large number of parameters with a cost approximately equal to the simulation of the physical problem.[9–11] Hence, for problems in which only a limited number of outputs are of interest, adjoint methods represent a powerful tool for efficiently obtaining gradient information. For this work, a discrete adjoint formulation is used to compute the gradient of objectives with respect to a wide variety of model and input parameters.

Using these gradients, a number of different methods may be employed to reduce the computational cost associated with Monte Carlo sampling. These methods typically rely on creating an inexpensive surrogate

*Graduate Student, Department of Mechanical Engineering; blockwoo@uwyo.edu, Student Member AIAA
†Postdoctoral Research Associate, Department of Mechanical Engineering; mrumpfke@uwyo.edu, Member AIAA
‡Postdoctoral Research Associate, Department of Mechanical Engineering; wyamazak@uwyo.edu, Member AIAA
§Professor, Department of Mechanical Engineering; mavripl@uwyo.edu, Associate Fellow AIAA

model for the simulation and range in complexity from utilizing simple extrapolations between sample points[5,6] to more sophisticated surrogate models, such as least square polynomials, multi-layer perceptron, radial basis functions (RBF), and Kriging. In particular, Kriging methods have gained popularity, especially in the field of computational fluid dynamics.[12–22] Due to the availability of efficient gradient evaluations, gradient enhanced Kriging methods have been developed and have demonstrated favorable results.[16,20,21] Once a surrogate of a simulation output is developed, the model may be sampled exhaustively to determine the output distribution and associated statistics.

In order to analyze hypersonic flows, numerous constitutive relations are required, each of which has a number of experimentally derived constants and parameters. These parameters are often the result of experimental measurements and have an associated variability. In general, this variability will have contributions from both aleatory and epistemic uncertainty. While aleatory uncertainty represents a variability due to inherent randomness in the system, epistemic uncertainty represents a lack of knowledge about the appropriate value to use for a quantity and is the dominant form of uncertainty encountered for hypersonic flows.[23,24] While a number of techniques exist for quantifying aleatory uncertainty, few are easily extended to epistemic uncertainty. In spite of this difficulty, regulatory agencies and design teams are increasingly being asked to specifically characterize and quantify epistemic uncertainty and separate its effect from that of aleatory uncertainty.[25] The goal of this work is to provide a framework in which both types of uncertainty can be quantified at reduced cost.

Within this paper, the use of gradient information for uncertainty quantification will be demonstrated for hypersonic flows. In Section II, details of the physical models and flow solver used for this work will be presented. In Section III, an outline of the discrete adjoint implementation is given and sample sensitivity results are given. Section IV provides details of the uncertainty quantification strategies proposed for this work as well as an overview of the surrogate models employed in this paper. Finally, Sections V and VI demonstrate the performance of the proposed uncertainty quantification strategies for a perfect gas and real gas test problem.

II. Flow Problem

In this section, the physical models and flow solver used in this paper will be outlined. For this work, the Navier-Stokes equations are solved numerically in two dimensions via a cell centered finite volume scheme on unstructured meshes using triangular and/or quadrilateral elements. In vector form, the Navier-Stokes equations are given by:

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \vec{F}(\mathbf{U}) = \nabla \cdot \vec{F}_v(\mathbf{U}) + \mathbf{S}(\mathbf{U}) \tag{1}$$

where $\mathbf{U}$ are the conserved flow variables, $\vec{F}$ is the inviscid flux, $\vec{F}_v$ is the viscous flux and S contains any source terms required for the physical model, such as reaction or energy coupling terms. Both perfect gas and non-equilibrium real gas physical models are examined for this work. The real gas model is the five species, two temperature model for non-ionizing air.[26] The Dunn-Kang model is used for the chemical kinetics required for the chemical non-equilibrium. The specific heats are calculated via 4th order polynomial curve fits covering various temperature ranges. The total enthalpy is calculated simply by integrating these curve fits and incorporating the proper heat of formation information.[27]

Two transport models have been examined for this work. The first model considered is based upon curve fits for species viscosity developed by Blottner *et al*[28] and use a mixing rule due to Wilke[29] to determine bulk quantities. Species diffusion coefficients are determined using the bulk transport quantities and the assumption of a constant Schmidt number across all species.[26] For the second model, viscosity, thermal conductivity and diffusion coefficients are calculated based on linear interpolation of collision integrals.[27,30]

The two phyiscal models considered in this work vary greatly in terms of complexity. For the perfect gas model, the only physical relations required are the perfect gas equation of state and Sutherland's law for viscosity. When freestream conditions are included, a perfect gas problem requires ten dimensional parameters to fully specify a flow problem. On the other hand, the real gas model has over 250 parameters embedded within the constitutive models for the reaction rates, transport coefficients, relaxation times and caloric equations of state.

In order to solve problems using the above two models, a two dimensional cell-centered finite volume code was written. The governing equations described above are first discretized in space and the solution is

advanced in time to steady state using a fully implicit approach. In semi-discrete form, the equations have the following form:

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{R}(\mathbf{U}) = 0 \tag{2}$$

The residual within each cell is given by the sum of the normal inviscid and viscous fluxes over all faces plus a cell centered contribution due to source terms. The inviscid flux is calculated using gradient reconstruction of primitive variables. The gradients are calculated using Green-Gauss contour integration over the cell. The limiter used within this code is a combination of a pressure switch and smooth Van Albada limiter, inspired by the experiences in references[26],[30] and.[31] The LDFSS flux function is used due to the ease with which it can be extended to additional equations. To extend this flux function to the real gas model, a frozen speed of sound is used.[31]

The result of the spatial discretization outlined above is a system of coupled ODE's which are solved implicitly using a BDF1 discretization. The result of this temporal discretization is a system of nonlinear equations which are solved using an inexact Newton's method. This inexact method employs a number of approximations to improve the performance of the nonlinear solver. Instead of solving the nonlinear system exactly, a fixed number of Newton iterations are performed. Additionally, instead of inverting the exact Jacobian, an approximate first-order Jacobian is used. Finally, in order to reduce the computational cost of each Netwon iteration, the preconditioner and transport quantities are frozen.

The solver described previously was validating using the standard test case of $5\,km/s$ flow over a circular cylinder with a super-catalytic, fixed temperature wall. The conditions for this test case can be found in Table 1. The results for this test case using the perfect gas and real gas model were compared against the well-validated codes LAURA and FUN2D.[32,33] The results for this benchmark are depicted in Figures 1 and 2.

Table 1. Benchmark flow conditions

$V_\infty =$	5 km/s
$\rho_\infty =$	0.001 kg/m^3
$T_\infty =$	200 K
$T_{wall} =$	500 K
$M_\infty =$	17.605
$Re_\infty =$	376,930
$Pr_\infty =$	0.72

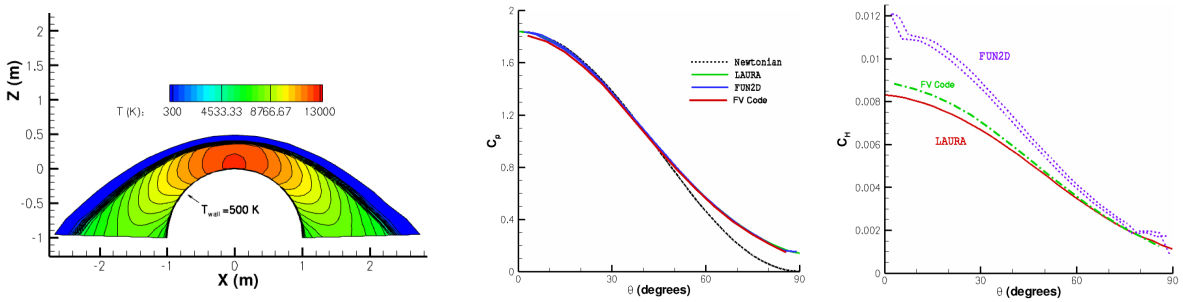

Figure 1. Validation of solver for $5\,km/s$ flow over circular cylinder using Perfect Gas Model. Left: Computed flow field temperature contours. Middle: Comparison of Surface Pressure Distribution with LAURA[34] and FUN2D [33] Right: Comparison of Surface Heating Distribution

III. Adjoint Sensitivity Analysis

With the flow solver detailed, an outline of the sensitivity procedure used to calculate the gradient of an objective is given. Further details of this adjoint implementation can be found in .[35] To determine the

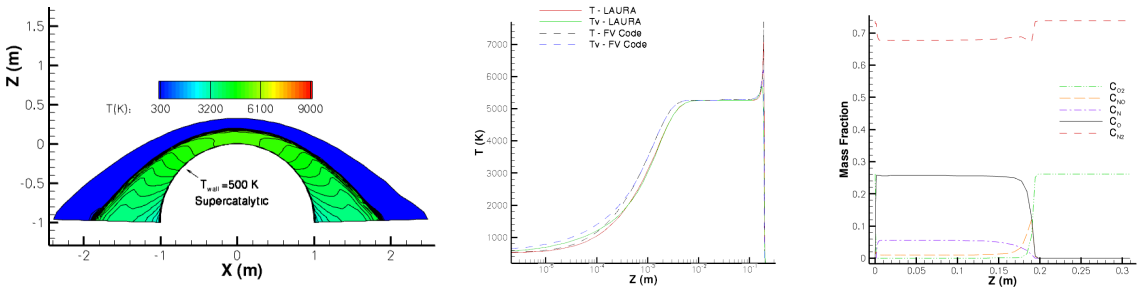

Figure 2. Validation of solver for $5\,km/s$ flow over circular cylinder. Left: Computed flow field temperature contours. Middle: Comparison of temperatures along centerline with LAURA[34] results running on equivalent mesh. Right: Species mass fractions along centerline.

gradient, the code is differentiated piece by piece and the final sensitivity is constructed using the chain rule. To illustrate this process, the following objective functional dependence is considered:

$$J \quad = J(D, \mathbf{U}(D)) \tag{3}$$

In addition to this objective, a constraint is needed. For the steady problems considered in this work, the constraint is that the spatial residual must equal zero.

$$\mathbf{R}(D, \mathbf{U}(D)) = 0 \tag{4}$$

Both the constraint and the residual have an explicit dependence on the input parameters, or design variables D, and an implicit dependence through the flow variables $\mathbf{U}$. In order to determine the sensitivity derivative, the objective can be differentiated using the chain rule as:[36]

$$\frac{dJ}{dD} = \frac{\partial J}{\partial D} + \frac{\partial J}{\partial \mathbf{U}} \frac{\partial \mathbf{U}}{\partial D} \tag{5}$$

The constraint may be differentiated in a similar manner. In this case, the derivative is equal to zero as the constraint must be satisfied for all admissible values of D and $\mathbf{U}$:

$$\frac{\partial \mathbf{R}}{\partial D} + \frac{\partial \mathbf{R}}{\partial \mathbf{U}} \frac{\partial \mathbf{U}}{\partial D} = 0 \tag{6}$$

Solving for $\frac{\partial \mathbf{U}}{\partial D}$ in the above equation and substituting into the objective derivative gives the forward sensitivity equation (also known as the tangent linear model).

$$\frac{dJ}{dD} = \frac{\partial J}{\partial D} - \frac{\partial J}{\partial \mathbf{U}} \frac{\partial \mathbf{R}}{\partial \mathbf{U}}^{-1} \frac{\partial \mathbf{R}}{\partial D} \tag{7}$$

As the equation shows, the residual Jacobian must be inverted once for each design variable. However, the same $\frac{\partial \mathbf{U}}{\partial D}$ may be used for each objective J. Due to the expense associated with inverting the residual Jacobian, the forward sensitivity approach is best suited for problems with few design variables and multiple objectives.

The adjoint sensitivity equation is found by taking the transpose of the forward equation.

$$\frac{dJ}{dD}^T = \frac{\partial J}{\partial D}^T - \frac{\partial \mathbf{R}}{\partial D}^T \frac{\partial \mathbf{R}}{\partial \mathbf{U}}^{-T} \frac{\partial J}{\partial \mathbf{U}}^T \tag{8}$$

where the last two terms can be replaced by the adjoint variable $\mathbf{\Lambda}$, defined as:

$$\frac{\partial \mathbf{R}}{\partial \mathbf{U}}^T \mathbf{\Lambda} = -\frac{\partial J}{\partial \mathbf{U}}^T \tag{9}$$

A sample adjoint solution for the $5km/s$ benchmark is found in Figure 3. This figure shows the adjoint variable for surface heating associated with the density. The magnitude of this variable roughly represents

the importance of the flow field on the objective of interest. As expected for surface heating, the adjoint variable is largest near the surface of the cylinder. Using this definition of the flow adjoint, the final sensitivity equation is given by:

$$\frac{dJ}{dD} = \frac{\partial J}{\partial D} + \mathbf{\Lambda}^T \frac{\partial \mathbf{R}}{\partial D} \tag{10}$$

Determining the solution of the flow adjoint equation roughly follows the procedure used to solve the analysis problem. A simplified preconditioner matrix is used to advance the adjoint solution in a defect-correction scheme.[36] The effect of the exact Jacobian required for the defect-correction solver is built up using automatically differentiated subroutines. The automatic differentiation used in this work is provided by the Tapenade Automatic Differentiation Engine.[37] Using this adjoint implementation, the sensitivity of an objective to any number of parameters can be computed with a constant amount of work.

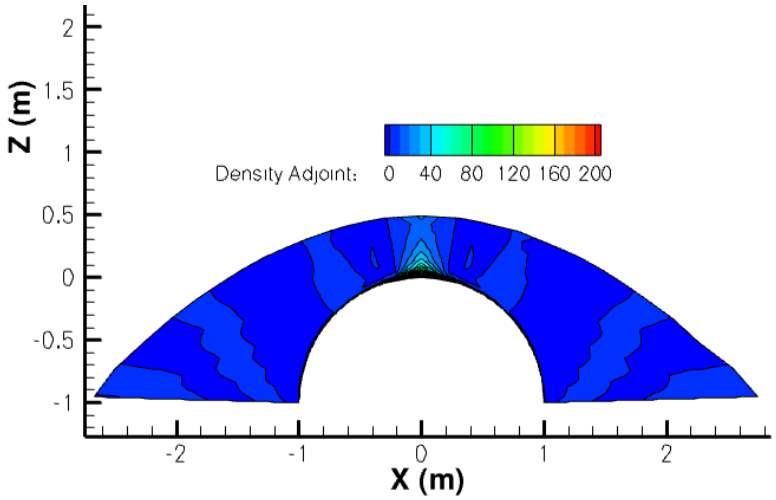

Figure 3. Density adjoint for integrated surface heating

To demonstrate the use of the adjoint, the sensitivities of surface heating with respect to parameters governing the chemical kinetics model and transport coefficients were calculated and are presented for the $5km/s$ real gas benchmark. The objective used for these demonstration results is integrated surface heating, given by the equation:

$$L = -\frac{\int_{\partial\Omega} k\nabla T \cdot \vec{n} + k_v \nabla T_v \cdot \vec{n} dA}{\frac{1}{2}\rho_\infty V_\infty^3} \tag{11}$$

The first variables examined relate to the specification of reaction rates. For the Dunn-Kang chemical kinetics model used within this work, the reaction rates take the following form:

$$K_f = C_f T_a^{\eta_f} e^{-\frac{E_{a,f}}{kT_a}} \tag{12}$$

$$K_b = C_b T_a^{\eta_b} e^{-\frac{E_{a,b}}{kT_a}} \tag{13}$$

where $E_{a,f}$ and $E_{a,b}$ represent the activation energy for the forward and backward reactions respectively, k is Boltzmann's constant and T_a is a characteristic temperature. The parameters examined in this case were C_f and C_b for each reaction, giving a total of 34 parameters. Figure 4 depicts the computed sensitivity of surface heating with respect to the forward and backward reaction rates using the adjoint procedure for the $5\,km/s$ benchmark case. It should be noted that due to the large discrepancy between the design variable and the objective, the sensitivity is expressed as fractional change in objective per fractional change in design variable (i.e. $\frac{\frac{dL}{L}}{\frac{dD}{D}}$).[35] As the results demonstrate, the reactions governing the production and breakdown of NO, as well as the oxygen recombination reactions, have the greatest influence on integrated surface heating.

In addition to reaction rates, the sensitivity with respect to parameters within the transport model was calculated. For the Blottner model, the curve fit of species viscosity is of the form:

$$\mu_s = 0.1e^{(A_s ln(T)+B_s)ln(T)+C_s} \tag{14}$$

Associated with each species are the parameters A_s,B_s, and C_s, giving a total of 15 parameters for the transport model. The sensitivity due to the transport coefficients is depicted in Figure 5. From these results, one can see that transport parameters seem to have a greater impact on surface heating, with the parameters associated with N_2 having the greatest individual contributions.

For the collision integral model, measured collision integrals between the five species at 2000 K and 4000 K are used and linear interpolation is used to determine the cross-section at the appropriate temperature.[27]

$$log_{10}(\Omega_{s,r}^{k,k}) = log_{10}(\Omega_{s,r}^{k,k})_{2000} + \left[log_{10}(\Omega_{s,r}^{k,k})_{4000} - log_{10}(\Omega_{s,r}^{k,k})_{2000}\right] \frac{ln(T) - ln(2000)}{ln(4000) - ln(2000)} \tag{15}$$

For the five species model, 15 independent collision interactions are possible. This fact gives a total of 60 parameters as two separate collision integrals ($\Omega_{s,r}^{1,1}$ and $\Omega_{s,r}^{2,2}$) are used at each temperature. In reality, it is likely that the uncertainty for the collision integrals measured at 2000K and 4000K are correlated. Due to this assumption, the sensitivities with respect to only the collision integrals at 2000 K were calculated. The results of this calculation are also presented in Figure 5. As the results show, the collisions involving N2 have the greatest effect on integrated surface heating. This result is unsurprising as N2 is the predominant species at the super-catalytic wall boundary condition.

As mentioned previously, the uncertainties associated with input and model parameters encountered within hypersonic problems are typically epistemic (i.e. the parameters are known to lie within a certain range but the distribution within the interval is unknown). Methods assuming aleatory uncertainties, such as polynomial chaos or simple Monte Carlo sampling, have likely underestimated the uncertainty associated with simulation objectives.[24,38] Although this work only provides a preliminary treatment of epistemic uncertainty, the framework presented in this paper based on gradient observations and surrogate models should provide a framework for exploring epistemic uncertainty quantification through either linear methods, optimization based approaches or augmented sampling strategies.[23,39]

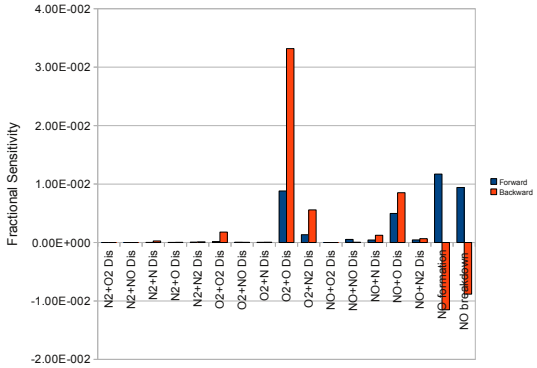

Figure 4. Sensitivity of surface heating with respect to forward and backward reaction rates.

IV. Uncertainty Quantification through Surrogate Models

The cornerstone of this approach is the construction of an inexpensive surrogate for the underlying design space associated with the simulation output. In particular, linear surrogates and Kriging models are considered. In order to reduce the computational cost associated with training the surrogate and provide additional information, derivatives with respect to the design variables were calculated using the discrete adjoint or the TLM depending on the dimension of the uncertainty space. A typical problem associated with uncertainty quantification is the so-called "curse of dimensionality", whereby the cost of quantifying uncertainty increases rapidly with the number of inputs. To address this problem, two different strategies

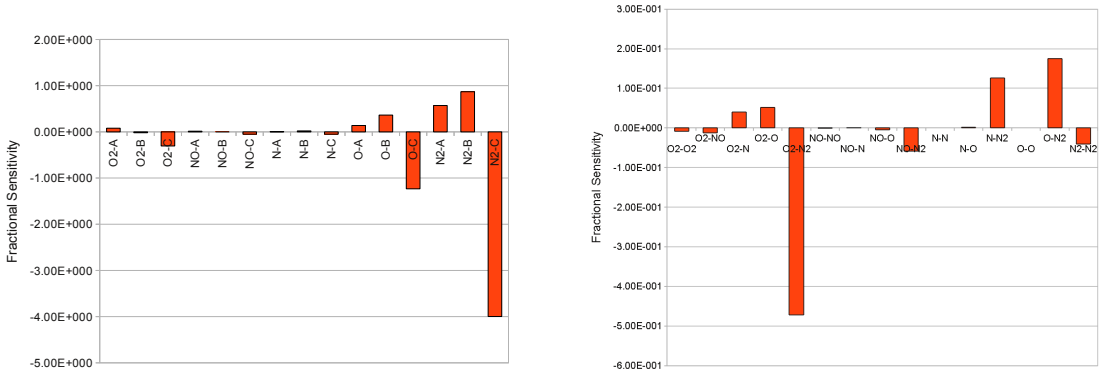

Figure 5. Left: Sensitivity of surface heating with respect to Blottner transport model parameters. Right: Sensitivity with respect to $\Omega_{s,r}^{1,1}$ at 2000 K

are proposed. First, only the uncertainty associated with input parameters greatly affecting the simulation output, determined through a sensitivity analysis, are considered, reducing the dimension of the problem at the outset. Second, the information gained at reduced cost through gradient observations will be exploited. For example, the function and gradient provide $M+1$ pieces of information for M inputs for the constant cost of roughly two function evaluations if adjoint techniques are utilized. Due to this additional information, the cost associated with representing the design space should be greatly reduced even as the dimension of the problem is increased.

IV.A. Local Sensitivity Approaches

In this section, the use of the first-order sensitivity derivatives for rapid uncertainty quantification is explored. As outlined previously, when a single simulation output is considered, the derivatives with respect to all input parameters can be computed using a single discrete adjoint solution. Using these derivatives, the variation of the objective throughout the design space is assumed to be linear. Hence, for these model, only a single function and gradient evaluation is necessary. For relatively small input uncertainties, the statistics based on these extrapolated values should approximate those based on the exact design space.

If one is only interested in the mean and standard deviation of an objective function, moment methods can provide an inexpensive way of approximating these values.[40,41] Moment methods are based on Taylor series expansions of the original nonlinear objective function $J(D)$ about the mean of the input D_0 given standard deviations σ_{D_j}. The resulting mean μ_J and standard deviation σ_J of the objective function are given to first-order (MM1) by:

$$\begin{aligned} \mu_J^{(1)} &= J(D_0) \\ \sigma_J^{(1)} &= \sqrt{\sum_{j=1}^{M} \left(\left. \frac{dJ}{dD_j} \right|_{D_0} \sigma_{D_j} \right)^2}, \end{aligned} \tag{16}$$

Because moment methods rely on an underlying distribution associated with the input parameters, these methods can only be used to propagate aleatory uncertainty. To propagate epistemic uncertainties, a similar equation based on interval addition may be used. Given the size of the input intervals Δ_{D_j}, the size of the output interval about the mean is given by:

$$\Delta_J^{(1)} = \sum_{j=1}^{M} \left| \left. \frac{dJ}{dD_j} \right|_{D_0} \Delta_{D_j} \right| \tag{17}$$

These linear models are inherently local in nature and therefore cannot account for the nonlinear nature of the underlying design space, limiting their usefulness to design variables with relatively small input

uncertainties or simple design spaces.[38] An additional drawback of the method of moments is that the probability density function of the output is not readily available. If the complete probability distribution function (PDF) of the objective is desired, Monte-Carlo sampling represents the most straight-forward approach. The number of simulations required for Monte-Carlo sampling can be dramatically reduced when sensitivity information is available by computing only a small number of simulations and extrapolating or interpolating the simulation results using the sensitivity information in between these simulation values.[5,6,42] Additionally, a global surrogate model for the function space can be created and sampled in place of the actual function. The simplest such surrogate model would be a linear representation of the function space in which the gradient is used to extrapolate function values about the mean. If more accuracy is required or the underlying function space is more complicated a better surrogate model, such as Kriging, is required.

IV.B. Linear Extrapolation

A useful application of the gradient is for extrapolation as discussed in Ghate and Giles.[6] The extrapolated function values can, for example, be used for an inexpensive Monte Carlo (IMC) simulation[5,43,44] for uncertainty analysis as extrapolation of the function value is much less expensive than a nonlinear function evaluation.

For the first-order sensitivities available from the discrete adjoint and tangent linear model, four different extrapolations may be used. These extrapolations include a linear extrapolation as well as three adjoint corrected extrapolations (see Giles *et al.*[45,46] for theoretical details on adjoint error correction). The linear extrapolation (Lin) represents a first-order Taylor series of the objective about the mean input parameters, given by:

$$J_{\text{Lin}} = J\Big(D_0, \mathbf{U}(D_0)\Big) + \left.\frac{dJ}{dD}\right|_{D_0} \cdot (D - D_0) \tag{18}$$

The adjoint corrected extrapolations rely on the fact that the objective function is a relatively inexpensive function for a given set of flow variables. Hence, the explicit dependence of the objective on the design variables may be treated exactly, linearizing only with respect to the flow variables. The adjoint corrected function evaluation of constant terms (ACCT) uses the product of the discrete adjoint and the residual to approximate the effect of changes in flow variables on the objective. The ACCT extrapolation is given by:

$$J_{\text{ACCT}} = J\Big(D, \mathbf{U}(D_0)\Big) + \mathbf{\Lambda}_{D_0}^T \cdot R\Big(D, \mathbf{U}(D_0)\Big),$$

Here, $R(D, \mathbf{U}(D))$ is the residual of the flow equations evaluated at the perturbed design points and $\mathbf{\Lambda}$ is the flow adjoint. The final two adjoint corrected extrapolations utilize information from the forward sensitivity, namely $\frac{\partial \mathbf{U}}{\partial D}$, to further approximate the effect of perturbations in the flow variables caused by a change in the design variables. The adjoint corrected linear extrapolation (ACLin) is given by:

$$J_{\text{ACLin}} \quad = \quad J_{\text{Lin}} + \mathbf{\Lambda}_{D_0}^T \cdot \mathbf{R}\Big(D, \mathbf{U}(D_0) + \left.\frac{d\mathbf{U}}{dD}\right|_{D_0} \cdot (D - D_0)\Big)$$

Finally, the adjoint corrected function evaluation of linearly extrapolated terms (ACLT) is given as:

$$\mathcal{J}_{\text{ACLT}} \quad = \quad \mathcal{J}\Big(D, q(D_0) + \left.\frac{dq}{dD}\right|_{D_0} \cdot (D - D_0)\Big) + \psi_{D_0}^T \cdot R\Big(D, q(D_0) + \left.\frac{dq}{dD}\right|_{D_0} \cdot (D - D_0)\Big) \tag{19}$$

Using these extrapolations, the objective throughout the design space can be approximated and used within an inexpensive Monte Carlo simulation. The most simple IMC approach would use a single set of function and derivative values evaluated at the mean quantities to represent the entire design space. Instead of taking samples from the simulation, function values are replaced with extrapolated values. For larger or more complicated uncertainty spaces, Dutch interpolation based on multiple samples may be used.[5,42,43]

For the hypersonic test cases presented in this paper, only results using linear extrapolation and ACCT are presented, although the infrastructure for all of the above extrapolation techniques is available. In order to demonstrate the above extrapolation techniques, the error encountered in extrapolating surface heating is plotted in Figure 6 for the $5km/s$ benchmark using the perfect gas model for one design variable, freestream

density. As the Figure demonstrates, both extrapolations work reasonably well for small perturbations ($< 5\%$) but break down as the size of the perturbation increases, with the error in the linear extrapolation increasing faster than ACCT. This behavior is expected as the ACCT extrapolation only linearizes with respect to flow variables and treats the explicit dependence of the objective on the design variables exactly.

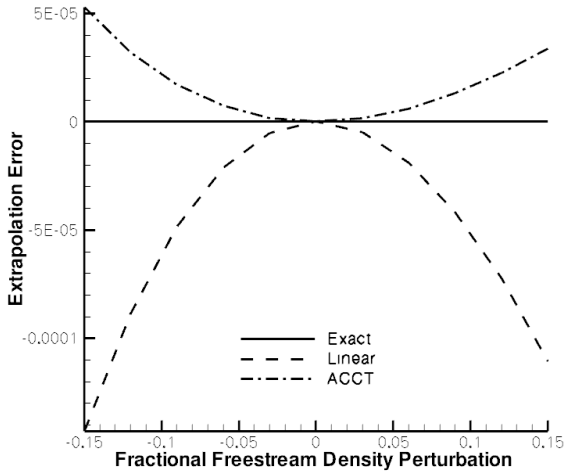

Figure 6. Error of Surface Heating Extrapolation for $5km/s$ test case using Linear and ACCT of Freestream Density

IV.C. Uncertainty Quantification Using Kriging Models

In order to provide a more accurate representation of the design space and propagate large input uncertainties, a more sophisticated representation of the function space is created using Kriging based surrogate models. Recently, the authors of this paper have developed gradient enhanced direct, as well as indirect, Kriging models for computational fluid dynamic simulations (also known as Co-Kriging models).[8,22,44] For these methods, the covariance of function/gradient and of gradients is considered in order to provide additional training points to the model at reduced computational cost. For each function/gradient evaluation, $M + 1$ pieces of information are obtained, where M is the dimension of the problem. This additional information can be used to greatly reduce the number of training points required to yield an accurate model. The construction of Kriging models is based on the assumption that the output variable J is a random variable obeying a Gaussian process, represented as:

$$J = N(m(D), K(D, D')) \tag{20}$$

where $m(D)$ is the mean function of the distribution and K is the covariance matrix of the training data. For gradient enhanced Kriging models, the covariance between function and derivative observations is considered, as these values obey the same Gaussian process due to the linearity of differentiation. For this work, the mean function is assumed constant, corresponding to ordinary Kriging. Using the observation data, represented by the vector $\mathbf{J}$, the behavior of J away from these observations can be predicted using the equation:

$$J^* = \mu_0 + k(D^*, D)K^{-1}\left(\mathbf{J} - \boldsymbol{\mu}_0\right) \tag{21}$$

For more details on the Kriging models used for this work, references [8,22,44] should be consulted.

Using the Kriging model, a number of uncertainty quantification strategies may be employed. For aleatory uncertainties, the surrogate model can be sampled using traditional Monte Carlo sampling at reduced cost. For epistemic uncertainties, Monte Carlo methods may also be employed, but the results can only be interpreted with regards to the interval of the output functional produced, with no inferred statistical distribution. Other approaches, such as Dempster-Shafer evidence theory also require large numbers of function evaluations.[47] In this case, the construction of a Kriging model remains an effective option for propagating epistemic uncertainties. In addition to sampling based approaches, epistemic uncertainty may also be propagated based on constrained optimization. This can either be performed by exhaustive sampling (capitalizing

on the reduced cost of the surrogate) or by optimization techniques such as the expected improvement (EI) techniques often invoked with Kriging methods. Using EI, the statistical distribution inherent in the Kriging model can be used to determine the areas in the design space with the highest probability of producing new extrema in the output.

V. Perfect Gas Demonstration Results

In order to demonstrate the use of derivative values for efficient uncertainty quantification, the $5\,km/s$ benchmark test using the perfect gas model was first examined. For these tests, the uncertain parameters are the freestream properties and the effect these parameters have on integrated surface heating is quantified. In particular, the freestream density, velocity, temperature, viscosity and thermal conductivities are treated as uncertain. Table 2 contains the mean values for these variables as well as the associated uncertainty. For this preliminary work, a 5% standard deviation in all variables except freestream velocity was assumed, inspired by the experiences of .[4,48] For freestream velocity, a standard deviation of $15.42m/s$ is used.[48] For these first tests, the uncertainties in the input parameters were assumed to be aleatory and distributed according to a normal distribution. Because this work ultimately relies on creating a surrogate model independent of input distribution, it can be extended to account for epistemic uncertainties. Preliminary Epistemic uncertainty results will be presented later in this section.

Table 2. Test Case flow conditions and Uncertainty for Perfect Gas Case

Variable	Mean Value	Standard Deviation
$V_\infty =$	5 km/s	15.42 m/s
$\rho_\infty =$	$0.001\ kg/m^3$	$5 \times 10^{-5}\ kg/m^3$
$T_\infty =$	200 K	10 K
$\mu_\infty =$	$1.3265 \times 10^{-5} kg/(m-s)$	$6.6325 \times 10^{-7} kg/(m-s)$
$k_\infty =$	$1.8576 \times 10^{-2} W/(m-K)$	$9.2880 \times 10^{-4} W/(m-K)$

In order to provide a baseline for comparison, Monte Carlo sampling of the flow solver was used to quantify the mean and standard deviation of the integrated surface heating based on the input distributions outlined above. In order to acquire the necessary statistics, 5000 simulation outputs were generated using Latin Hypercube sampling of the input parameters. The convergence of the mean and standard deviation for integrated surface heating are found in Figure 7. Table 3 contains the fully converged statistics of the Monte Carlo samples. These statistics will serve as the metric by which the other uncertainty quantification techniques are compared.

Table 3. Monte Carlo Statistics for Surface Heating after 5000 Samples for Perfect Gas Case

Mean (μ)	1.47069×10^{-2}
Standard Deviation (σ)	5.38407×10^{-4}
95% Confidence Interval	$\pm 7.3218\%$

V.A. Linear Results

As a first attempt at quantifying the uncertainty present in the $5km/s$ benchmark, a first-order moment method is used to estimate the mean and standard deviation of the output. These results are found in Table 4. As the table shows, the moment method slightly underestimates the mean and standard deviation. Remarkably, the results are surprisingly good considering the highly nonlinear nature of the underlying physical problem. It should be noted that the input uncertainties specified for this test problem are relatively small and it is likely the results would not agree as well for larger input uncertainties .[3]

In order to demonstrate the utility of linear extrapolation for the purposes of IMC sampling, a single function and gradient evaluation is performed and linear extrapolation is used to determine the function value

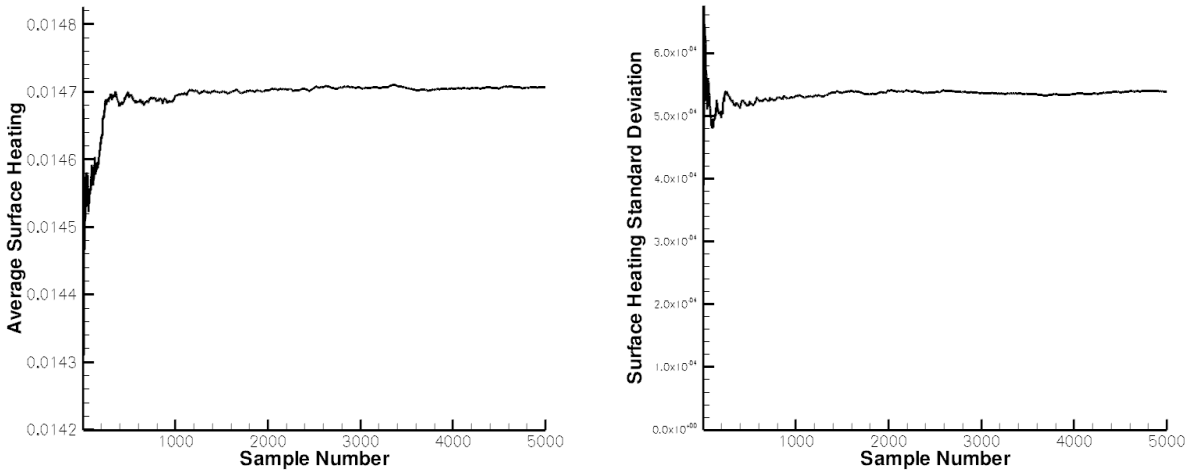

Figure 7. Monte Carlo convergence of Surface Heating for Average and Variance for Perfect Gas Case.

Table 4. First-Order Moment Method Statistics for Surface Heating for Perfect Gas Case.

	Moment Method	Nonlinear Monte Carlo
Mean (μ)	1.46917×10^{-2}	1.47069×10^{-2}
Standard Deviation (σ)	5.32578×10^{-4}	5.38407×10^{-4}
95% Confidence Interval	$\pm 7.2501\%$	$\pm 7.3218\%$

about the mean. In principle, all the extrapolation methods presented in the previous subsection can be applied. In practice however, the adjoint corrected approaches have difficulties with the evaluation of the flow residual for large perturbations. In order to eliminate any source of bias, the same 5000 samples locations used for the nonlinear Monte Carlo sampling are used for the extrapolation. In addition to calculating statistics based on a linear representation of the design space, the output probability distribution function is also calculated and compared to the distribution found via Monte Carlo sampling. Table 5 contains the statistics for the linear extrapolation model. The statistics produced based on linear extrapolation are nearly equal to those produced by the method of moments. Although the linear extrapolation approach does produce statistics slightly closer to those from the nonlinear Monte Carlo sampling, this result is likely due to the limited number of samples used to build up the output distribution. As the number of samples is increased, the linear extrapolation statistics approach the method of moments results as expected. The distribution of the output for the linear extrapolation model is plotted and compared to the nonlinear Monte Carlo results in Figure 8. As the figure shows, the distribution near the mean is predicted well but some details begin to be lost at approximately one standard deviation from the mean.

Table 5. IMC Results using Linear Extrapolation for Surface Heating for Perfect Gas Case.

	Linear Extrapolation	Nonlinear Monte Carlo
Mean (μ)	1.46970×10^{-2}	1.47069×10^{-2}
Standard Deviation (σ)	5.35082×10^{-4}	5.38407×10^{-4}
95% Confidence Interval	$\pm 7.2815\%$	$\pm 7.3218\%$

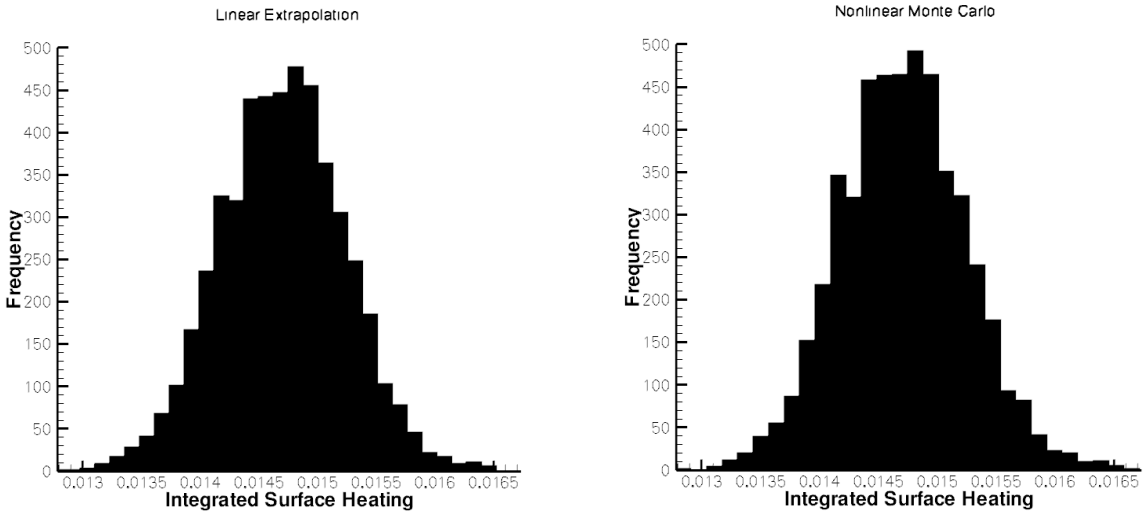

Figure 8. Histogram of Surface Heating from Linear Extrapolation compared to Nonlinear Monte Carlo results

V.B. Kriging Results

In order to test the applicability of Kriging models for uncertainty quantification and improve upon the previous linear results, Kriging models with varying numbers of training points were constructed and the output statistics based on these models were computed. Table 6 shows the mean and standard deviation from an IMC simulation using Kriging models created exclusively from function evaluations. As the table shows, approximately 40 function evaluations are required to match the nonlinear Monte Carlo statistics. In addition to statistics, the Kriging model can be used to produce a PDF of the output. Figure 9 contains the approximate PDF for a Kriging model based on 40 function evaluations. As the figure shows, although the peak of the PDF is not matched as well as the linear extrapolation PDF, the overall shape of the curve is captured over a wider range of the distribution.

Table 6. IMC Results using Kriging Model for Surface Heating

Sample Points	Mean	Standard Deviation	95% Confidence Interval
10	1.47550×10^{-2}	5.48708×10^{-4}	$\pm 7.4376\%$
20	1.47090×10^{-2}	5.39907×10^{-4}	$\pm 7.3412\%$
30	1.47150×10^{-2}	5.46086×10^{-4}	$\pm 7.4222\%$
40	1.47100×10^{-2}	5.38433×10^{-4}	$\pm 7.3206\%$
50	1.47070×10^{-2}	5.38767×10^{-4}	$\pm 7.3267\%$
60	1.47060×10^{-2}	5.38266×10^{-4}	$\pm 7.3204\%$

In order to reduce the number of samples required to accurately represent the function space, a gradient enhanced model is also examined. Table 7 contains the statistics using a direct Co-Kriging model. As the table shows, far fewer function/gradient evaluations are required to match the nonlinear Monte Carlo statistics. With only 4 function/gradient evaluations, the statistics are matched quite well. Considering that gradient evaluations are approximately the same cost as function evaluations, this result represents a large improvement over the 40 function evaluations required for the function-only Kriging model. In addition to capturing the output statistics, Figure 10 shows the ability of the Co-Kriging model to capture the PDF of the output with relatively few function/gradient evaluations.

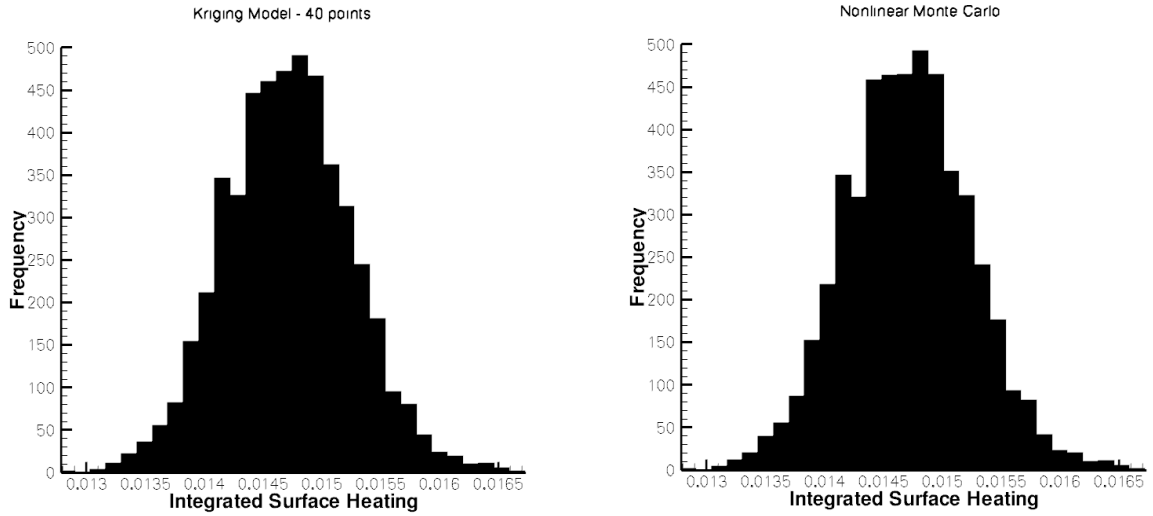

Figure 9. Histogram of Surface Heating from Kriging Model compared to Nonlinear Monte Carlo results

Table 7. IMC Results using Kriging Model for Surface Heating

Sample Points	Mean	Standard Deviation	95% Confidence Interval
4	1.47070×10^{-2}	5.38544×10^{-4}	$\pm 7.3236\%$
8	1.47090×10^{-2}	5.39268×10^{-4}	$\pm 7.3325\%$
10	1.47070×10^{-2}	5.38544×10^{-4}	$\pm 7.3236\%$
50	1.47070×10^{-2}	5.38424×10^{-4}	$\pm 7.3220\%$

V.C. Epistemic Uncertainty Results

In order to demonstrate the utility of gradient information for propagating epistemic uncertainty, the above perfect gas problem was re-examined with the input uncertainties treated as epistemic instead of aleatory. For this problem, gradient information is leveraged in two different ways. As a first step, Equation 17 is used to estimate the output interval by assuming the objective behaves linearly. In order to provide a comparison for this approach and further demonstrate the manner in which gradients may be used, a constrained optimization problem was solved in order to determine the associated output interval based on given input intervals. For this optimization, the L-BFGS algorithm was used due to the availability of gradient observations.

In order to consider the propagation of epistemic uncertainty, the previously defined problem must be recast and the input parameters must be defined in the interval sense. Within the literature, uncertain parameters are typically reported as the mean value plus/minus an interval width. For the aleatory case, this reported interval width was assumed to represent a 95% confidence interval for the parameter. For the epistemic case, two cases were considered. First, the reported uncertainty was assumed to represent the half-width of the input interval (i.e. the upper bound is the mean plus the reported uncertainty and the lower bound is the mean minus the reported uncertainty). This interpretation is the most likely meaning from the reported values. In order to provide a more difficult test for the method, a four sigma half-width interval was also used. If the reported values were indeed aleatory in nature, then taking an interval of four sigma on either side of the mean would ensure that virtually all parameter values are included in the analysis. Although this is not a likely interpretation of the data, it was considered to provide a more difficult test problem for the linear analysis and optimization. The associated input intervals for each scenario are given in Table 8.

As a first attempt to propagate epistemic uncertainty for the perfect gas problem, local sensitivity derivatives were used to estimate the output interval of the objective using Equation 17. For this linearized ap-

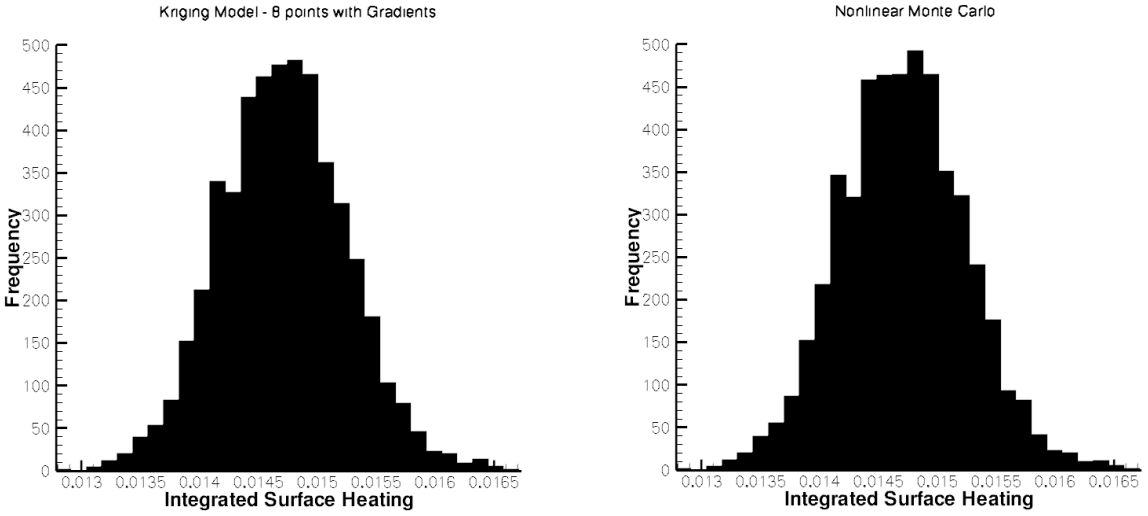

Figure 10. Histogram of Surface Heating from Gradient Enhanced Kriging Model compared to Nonlinear Monte Carlo results

Table 8. Test Case Input Intervals for Perfect Gas Case

Variable	Lower Bound (2 σ)	Upper Bound (2 σ)	Lower Bound (4 σ)	Upper Bound (4 σ)
$V_\infty(m/s) =$	4969.16	5030.84	4938.32	5061.68
ρ_∞ (kg/m^3)=	0.0009	0.0011	0.0008	0.0012
$T_\infty(K) =$	180	220	160	240
$\mu_\infty(kg/(m-s)) =$	1.19385×10^{-5}	1.45915×10^{-5}	1.0612×10^{-5}	1.5918×10^{-5}
$k_\infty(W/(m-K)) =$	1.6718×10^{-2}	2.0434×10^{-2}	1.4861×10^{-2}	2.2291×10^{-2}

proach, both input intervals were considered. Table 9 contains the results for this linearized analysis. The percentage uncertainty for this case is interval width divided by the average value. The output interval is assumed to be centered about the mean value and the upper and lower bounds are determined using Equation 17 and the half-width of each of the input intervals.

Table 9. Perfect Gas Output interval based on Linear Method

Input Interval	Lower Bound	Upper Bound	Interval Width
2 σ	1.3121×10^{-2}	1.6262×10^{-2}	21.38%
4 σ	1.1550×10^{-2}	1.7833×10^{-2}	42.76%

In order to provide a comparison for the linearized results, an optimization approach is used to determine the minimum and maximum possible objective values given the bounds on the input parameters. Because of the simplicity of the design space for this problem, the optimization results will be assumed to be correct results. For a more complicated design space, this assumption may be poor as the optimization may become stuck in local extrema and the optimization results may need to be further validated by an exhaustive sampling approach .[23] Figure 11 contains the results for both the minimization and maximization for both input intervals. As the Figure shows, the optimization rapidly converges to both extrema. For this simple problem, these extrema simply occur at the extremes of the input intervals. Table 10 contains the resulting output intervals for each input interval. Comparing the two methods, it can be seen that the linearized approach produces nearly identical output intervals when compared to the optimization approach. Based on the other results presented for this problem, this fact should not be surprising as the linear models appear

to be sufficient for representing the design space of the perfect gas model.

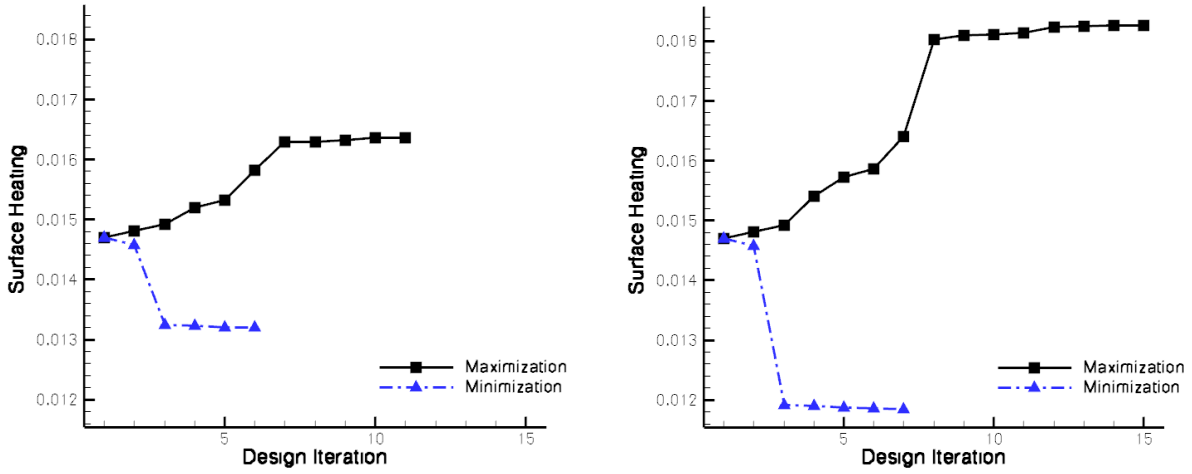

Figure 11. Optimization results for surface heating based on Left: the 2σ equivalent interval and Right: the 4σ equivalent interval.

Table 10. Perfect Gas Output interval based on Optimization Method

Input Interval	Lower Bound	Upper Bound	Interval Width
2 σ	1.3201×10^{-2}	1.6359×10^{-2}	21.49%
4 σ	1.1844×10^{-2}	1.8262×10^{-2}	43.68%

In order to demonstrate the error encountered by treating epistemic uncertainties as aleatory, the above results can be compared to the corresponding interval based on the aleatory results. In order to convert the aleatory results to an interval, the same procedure used on the input parameters can be used. Hence, a four sigma or eight sigma equivalent interval on the output is used for the respective input intervals. These results are compared to the epistemic results produced by the optimization in Table 11. As the table shows, treating these parameters as aleatory greatly underestimates the uncertainty in the output if the parameters are in reality epistemic. For this particular problem, the uncertainty assuming aleatory uncertainty in the inputs is approximately 75% of the variability present when the inputs are assumed to have epistemic uncertainty. Hence, methods which can only account for aleatory uncertainties likely underestimate the output variability in the presence of epistemic uncertainty.

Table 11. Comparison to Alleatory for Perfect Gas case

Input Interval	Epistemic Interval Width	Equivalent Alleatory Width
2 σ	21.49%	14.64%
4 σ	43.68%	29.29%

VI. Real Gas Results

In order to provide a more difficult test of the proposed uncertainty quantification strategies, the $5km/s$ benchmark case was reconsidered using the real gas model. As stated previously, the real gas model contains approximately 250 parameters. In order to reduce the computational cost associated with validating the uncertainty quantification and constructing the Kriging surface, only fifteen parameters are treated as uncertain. These fifteen parameters were chosen based on similar uncertainty studies found in previous

work[4,48] as well as the sensitivity analysis presented in Section III. The parameters chosen for this study, as well as the assumed standard deviations for these variables, can be found in Table 12. The choice of freestream parameters was influenced by reference,[48] while the particular choice of cross-sections was based on reference.[4] These choices account for ten variables. Finally, based on the sensitivity analysis presented in Section III, the five most important reaction terms were chosen in order to account for uncertainty with respect to more varied parameters. The uncertainty for these parameters was taken from.[49] If this work was being performed in isolation, the sensitivity analysis alone would suffice for choosing the parameters. For the most part, the results of the sensitivity analysis are in line with the choice of variables based on[4,48] with only slight differences with regard to the choice of cross-sections.

Table 12. Real Gas Model Parameters

Number	Variable	Standard Deviations
1	$V_\infty(m/s)$	15.42
2	$\rho_\infty\ (kg/m^3)$	5%
3-4	$\Omega^{1,1}_{N2-N2}, \Omega^{2,2}_{N2-N2}$	10%
5-6	$\Omega^{1,1}_{N2-N}, \Omega^{2,2}_{N2-N}$	10%
7-8	$\Omega^{1,1}_{N2-O}, \Omega^{2,2}_{N2-O}$	10%
9-10	$\Omega^{1,1}_{N2-O2}, \Omega^{2,2}_{N2-O2}$	10%
11	$log_{10}(C_{b-O2+O \leftrightharpoons 2O+O})$	0.5
12	$log_{10}(C_{f-N2+O \leftrightharpoons NO+N})$	0.5
13	$log_{10}(C_{b-N2+O \leftrightharpoons NO+N})$	0.5
14	$log_{10}(C_{f-NO+O \leftrightharpoons O2+N})$	0.5
15	$log_{10}(C_{b-NO+O \leftrightharpoons O2+N})$	0.5

For the collision integrals, the mean values at 2000 K and 4000 K were perturbed by the same amount, shifting the collision integral curve used for interpolation. This treatment was chosen due to the assumption that the uncertainty in these collision integral values is correlated.

In order to provide a validation metric for the linear and surrogate based results, Monte Carlo sampling based on Latin hypercube was used to determine the average and variance of the surface heating. For this test case, the input parameters were assumed to be aleatory and governed by a normal distribution. As was the case with the perfect gas test case, the reported uncertainties were assumed to represent a 95% confidence interval. Due to the limited computational budget, only 475 samples were possible. The convergence of the average and variance for the sampling is found in Figure 12. As the plot demonstrates, the average and variance appear to be converged despite the limited number of sample points. In spite of the appearance of convergence, error bounds should be established for the average and variance in order to provide a means of comparison with the alternative UQ strategies proposed.

For Monte Carlo sampling, the convergence of function statistics obeys a normal distribution with a variance dependent on the number of samples. For the mean, the standard deviation for this convergence distribution is given by:[50]

$$\sigma_N = \frac{\sigma}{\sqrt{N}} \tag{22}$$

where σ is the underlying standard deviation of the distribution being sampled (that is the objective function, surface heating). This standard deviation is in turn estimated by the sample variance (S^2) of the Monte Carlo results. Using this standard deviation estimate, a 95% confidence interval for the average prediction can be established based on a two-σ bound.

To provide bounds for the sample variance, its variance most also be estimated. For i.i.d samples, the variance of the sample variance can be calculated as:[51]

$$Var(S^2) = \sigma^4 \left(\frac{2}{N-1} + \frac{K}{N} \right) \tag{23}$$

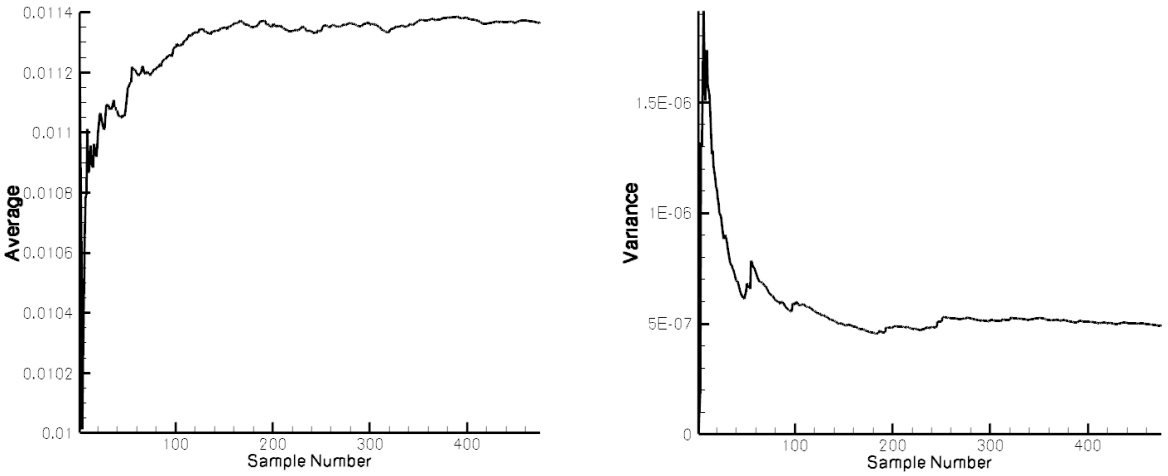

Figure 12. Monte Carlo convergence of Surface Heating for Average and Variance for Real Gas Case.

where K is the kurtosis of the distribution being sampled. The kurtosis can be estimated using the sample kurtosis G, given as:[52]

$$G = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2\right)^2} - 3 \tag{24}$$

Table 13 contains the mean and standard deviation predictions based on the 475 sample points. Along with these results are associated 95% confidence intervals for each of these predictions. These results represent the standard by which the other uncertainty quantification strategies will be judged.

Table 13. Real Gas Monte Carlo Results

Statistic	Value	Error Bound	95% Lower Bound	95% Upper Bound
Average	1.1368×10^{-2}	$\pm 6.46 \times 10^{-2}$	1.1303×10^{-2}	1.1433×10^{-2}
Variance	4.9562×10^{-7}	$\pm 5.10 \times 10^{-8}$	4.4465×10^{-7}	5.4660×10^{-7}
Standard Deviation	7.0400×10^{-4}	$\pm 3.63 \times 10^{-5}$	6.6682×10^{-4}	7.3932×10^{-4}
95% Confidence Interval	$\pm 12.385\%$	$\pm 0.637\%$	$\pm 11.7315\%$	$\pm 13.0071\%$

VI.A. Linear Results

As a first attempt to quantify the uncertainty for the real gas model, local sensitivity derivative values were calculated and used as a means of ranking the individual contributions to uncertainty made by each variable as well as estimate the total uncertainty in the surface heating due to the 15 variables. In order to assess the individual contributions to uncertainty by each variable, the sensitivity derivative for each variable was multiplied by its associated standard deviation. The magnitudes for these contributions are plotted in Figures 13. From these results, freestream density and the Σ_{N_2-O} collision integral contribute the most to the surface heating uncertainty. It should be noted that these represent contributions to non-dimensional surface heating. Because freestream density and velocity are used to non-dimensionalize surface heating, the relative importance of the variables is altered when dimensional surface heating is considered. Figure 13 shows the relative contributions to dimensional surface heating. When the dimensional objective is analyzed, freestream density and velocity become the biggest contributors to the uncertainty. These results are in agreement with other uncertainty analysis performed in other works.[4,48] The advantage of this approach being that this information was acquired from a single flow solution and adjoint solution.

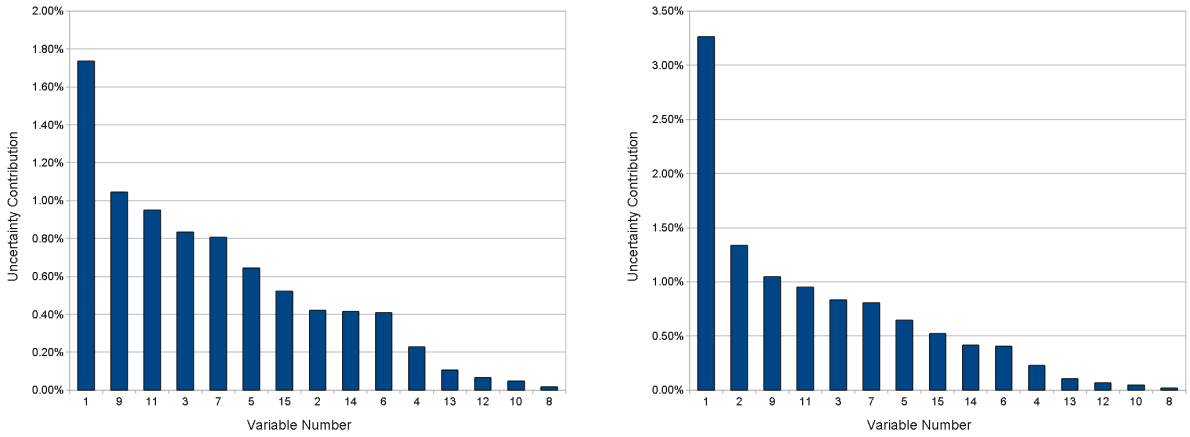

Figure 13. Left: Contribution of each variable to uncertainty of non-dimensional integrated surface heat using linearized analysis. Right: Contribution of each variable to uncertainty of dimensional integrated surface heat using linearized analysis.

The total uncertainty in surface heating is estimated using Equation 16. The statistics from this localized treatment are compared to the Monte Carlo results in Table 14. As these results demonstrate, the linearized approach greatly underestimates the confidence interval on the surface heating. Despite underestimating the total confidence interval for surface heating, the linearized estimate is likely accurate enough to give a rough approximation of the quality of an answer and the relative magnitudes of the individual uncertainty contributions appear to agree with previous uncertainty studies.[4,48]

Table 14. First-Order Moment Method Statistics for Surface Heating for Real Gas Case.

	Moment Method	Nonlinear Monte Carlo
Mean (μ)	1.1238×10^{-2}	1.1368×10^{-2}
Standard Deviation (σ)	3.1033×10^{-4}	7.0400×10^{-4}
95% Confidence Interval	$\pm 5.5231\%$	$\pm 12.385\%$

VI.B. Kriging Results

In order to improve upon the results of the linearized analysis, the use of Kriging models for representing the real gas design space was investigated. Using function values only, Kriging models using varying numbers of training points were created and the statistics based on these models were calculated. The convergence of the average and the 95% confidence interval are plotted in Figure 14 as a function of training points. As the plot demonstrates, the average rapidly converges to the interval corresponding to the Monte Carlo results. The confidence interval prediction is essentially converged after approximately 100 training points. The confidence interval predicted by the Kriging surface is relatively low compared to the Monte Carlo results. However, the Kriging model using a fifth of the number of samples used for the Monte Carlo results gives a relatively good estimate of the surface heating uncertainty.

Due to robustness issues in calculating the sensitivity derivatives for design variables away from the mean, Co-Kriging results using the real gas model are not available. Because the calculation of derivatives requires the inversion of the residual Jacobian (or transpose of the Jacobian in the case of the adjoint), the calculation of sensitivity derivatives is affected greatly by the condition number of the Jacobian and the starting point of the iterative solver. For small departures from the mean, the Jacobian inversion can be performed successfully. As design variables farther from the mean are considered, the Jacobian requires an impractical amount of diagonal padding to successfully calculate the adjoint. It should be noted that these same areas of the design space presented problems for the flow solution. However, in the case of the flow solution, physics based robustness enhancers, such as freezing transport terms or adjusting limiter constants, may be employed. To improve the robustness of the adjoint solution process, additional preconditioning techniques

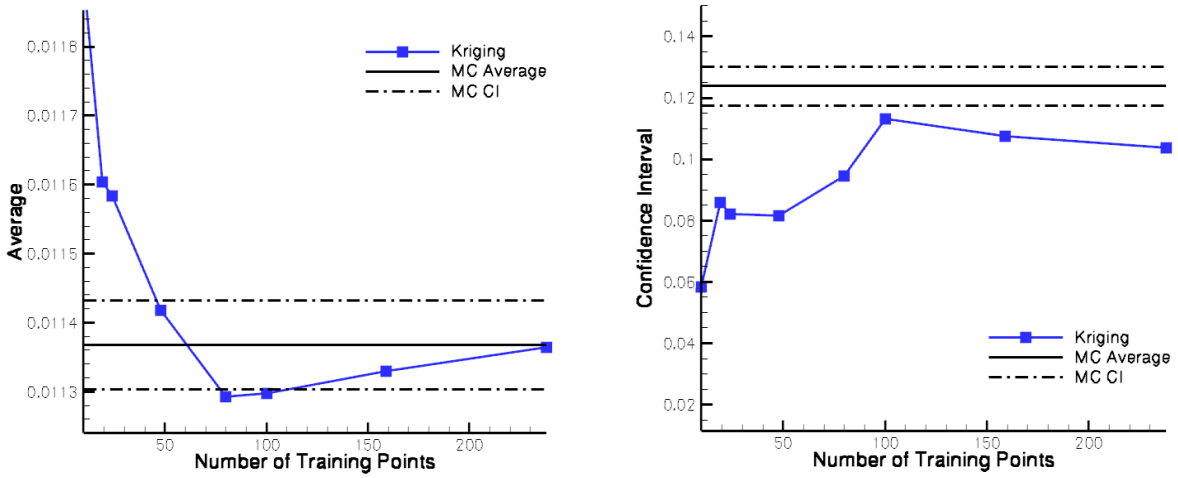

Figure 14. Left: Average prediction of Kriging model as function of training points. Right: Confidence Interval prediction of Kriging model as function of training points.

as well as more sophisticated matrix inversion algorithms should be examined; however, this examination is left for future work.

VII. Conclusions and Future Work

Within this paper, the use of gradient observations in conjunction with inexpensive surrogate models for uncertainty quantification in hypersonic flows has been demonstrated. Using sensitivity derivatives calculated from the discrete adjoint, input parameter uncertainty was propagated to the output using the method of moments and inexpensive Monte Carlo based on linear extrapolation. To provide a more sophisticated model of the design space, Kriging and Co-Kriging models were also examined as a basis for inexpensive Monte Carlo. To demonstrate the proposed methods, two uncertainty problems were considered. In order to test a wide variety of models on a relatively simple problem, the uncertainty of surface heating associated with the freestream conditions was examined based on the perfect gas model. Using the method of moments and an inexpensive Monte Carlo based on linear extrapolation, the mean and standard deviation of the integrated surface heating was estimated. These results compared favorably to the statistics derived from Monte Carlo sampling based on the nonlinear flow solver. In addition to the mean and standard deviation of the output, the probability distribution function (PDF) for the surface heating was approximated based on the linear extrapolation results and compared to the PDF from the nonlinear Monte Carlo results. These first-order methods provided reasonable estimates of the output statistics but were slightly optimistic when compared to the results of the nonlinear Monte Carlo sampling. In an attempt to improve the agreement of the statistics and approximate PDF's from the inexpensive Monte Carlo procedure, linear extrapolation was replaced by a Kriging model. For this method, a Kriging model of the function space was created and used to determine the function value at the perturbed design variables. Using relatively few function samples (approximately 40 for 5 design variables), the Kriging-based IMC accurately matched the statistics of the nonlinear Monte Carlo results and accurately captured many of the features of the output PDF, especially toward the tails of the distribution. Additionally, a gradient enhanced Kriging model was created. By using gradient information, the required number of function/gradient evaluations needed to accurately match the output statistics was greatly reduced.

In order to provide a more difficult test for the proposed uncertainty quantification strategies, the uncertainty of integrated surface heating associated with parameters within the real gas model was quantified. In order to provide a rough estimate of the total uncertainty and analyze the contribution to uncertainty by each variable, a linear model based on sensitivity derivatives was used. Although the total uncertainty estimate was greatly underestimated relative to the nonlinear Monte Carlo results compiled for this problem, this linearized analysis was able to provide an estimate of which variables provide the greatest contribution

to the overall uncertainty. The variable ranking based on this estimate was in general agreement with the individual contribution estimates of previous uncertainty studies.[4,48] To provide a better estimate of the overall uncertainty in surface heating, a series of Kriging models were created based on function evaluations. Using 100 function evaluations, the Kriging model was able to closely replicate the statistics of the full Monte Carlo sampling, representing nearly a factor of five savings.

In order to improve upon these results and further reduce the associated cost of uncertainty, the use of gradient-enhanced Kriging models for the larger dimensional real gas model should be further explored. Additionally, the propagation of epistemic uncertainty using the constrained optimization approach should be considered for the more difficult real gas problem. For both of these goals, significant robustness enhancements in the discrete adjoint solver must be made. Once these robustness enhancements have been made, further work will explore dimension reduction techniques as well as multiple fidelity methods within the context of real gas problems and Kriging models.

Acknowledgments

This work is supported by a DOE Computational Science Graduate Fellowship (Brian Lockwood) and the US Air Force Office of Scientific Research under AFOSR Grant number FA9550-07-1-0164.

References

[1]Luckring, J. M., Hemsch, M. J., and Morrison, J. H., "Uncertainty in computational aerodynamics," *41st AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January 2003, AIAA Paper, 2003-0409.

[2]Gumbert, C. R., Newman, P. A., and Hou, G. J., "Effect of random geometric uncertainty on the computational design of 3-D wing," *20th AIAA Applied Aerodynamics Conference*, St. Louis, MO, June 2002, AIAA Paper, 2002-2806.

[3]Michael J. Wright, D. B. and Chen, Y.-K., "Probabilistic Modeling of Aerothermal and Thermal Protection Material Response Uncertainties," *AIAA Journal*, Vol. 45, No. 2, February 2007.

[4]Palmer, G. E., "Uncertainty Analysis of CEV LEO and Lunar Return Entries," *39th AIAA Thermophysics Conference*, Miami, FL, June 2007, AIAA Paper, 2007-4253.

[5]Ghate, D. and Giles, M. B., "Inexpensive Monte Carlo uncertainty analysis," *Recent Trends in Aerospace Design and Optimization, Tata McGraw-Hill, New Delhi*, 2006, pp. 203–210.

[6]Ghate, D. P. and Giles, M. B., "Efficient Hessian Calculation using Automatic Differentiation," *25th AIAA Applied Aerodynamics Conference*, Miami, FL, June 2007, AIAA Paper, 2007-4059.

[7]Rumpfkeil, M. P. and Mavriplis, D. J., "Efficient Hessian Calculations using Automatic Differentiation and the Adjoint Method," AIAA Paper, 2010-1268, January, 2010.

[8]Yamazaki, W., Rumpfkeil, M. P., and Mavriplis, D. J., "Design Optimization Utilizing Gradient/Hessian Enhanced Surrogate Model," *40th Fluid Dynamics Conference and Exhibit*, Chicago, IL, June 2010, AIAA Paper, 2010-4363.

[9]Pironneau, O., "On Optimum Design in Fluid Mechanics," *Journal of Fluid Mechanics*, Vol. 64, No. 1, 1974, pp. 97–110.

[10]Jameson, A., "Optimum Aerodynamic Design Using Control Theory," *Computational Fluid Dynamics Review, Hafez, M., Oshima, K. (eds)*, Wiley: New York, pp. 495-528, 1995.

[11]Errico, R. M., "What is an adjoint model?" *Bulletin of the American Meteorological Society*, Vol. 8(11), 1997, pp. 25772591.

[12]Cressie, N., "The Origins of Kriging," *Mathematical Geology*, Vol. Vol. 22, No. 3, 1990, pp. 239–252.

[13]Koehler, J. R. and Owen, A. B., "Computer Experiments," *Handbook of Statistics, Ghosh, S., Rao, C.R., (Eds.)*, pp. 261-308, 1996.

[14]Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. Vol. 13, 1998, pp. 455492.

[15]Simpson, T. W., Korte, J. J., Mauery, T. M., and Mistree, F., "Comparison of Response Surface and Kriging Models for Multidisciplinary Design Optimization," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, 1998, AIAA Paper, 98-4758.

[16]Chung, H. S. and Alonso, J. J., "Using Gradients to Construct Cokriging Approximation Models for High-Dimensional Design Optimization Problems," *40th AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, January 2002, AIAA Paper, 20020317.

[17]Martin, J. D. and Simpson, T. W., "Use of Kriging Models to Approximate Deterministic Computer Models," *AIAA Journal*, Vol. Vol. 43, No.4, 2005, pp. 853–863,.

[18]Jeong, S., Murayama, M., and Yamamoto, K., "Efficient Optimization Design Method Using Kriging Model," *Journal of Aircraft*, Vol. Vol. 42, No. 2, 2005, pp. 413–420.

[19]Peter, J. and Marcelet, M., "Comparison of Surrogate Models for Turbomachinery Design," *WSEAS Transactions on Fluid Mechanics*, Vol. Vol. 3(1), 2008.

[20]Laurenceau, J. and Sagaut, P., "Building Efficient Response Surfaces of Aerodynamic Functions with Kriging and Cokriging," *AIAA Journal*, Vol. Vol. 46, No. 2, 2008, pp. 498–507.

[21]Laurenceau, J. and Meaux, M., "Comparison of Gradient and Response Surface Based Optimization Frameworks Using Adjoint Method," AIAA Paper, 2008-1889, 2008.

[22]Yamazaki, W., Mouton, S., and Carrier, G., "Efficient Design Optimization by Physics-Based Direct Manipulation Free-Form Deformation," *12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Victoria, Canada, September 2008, AIAA Paper, 2008-5953.

[23]Helton, J. C., Johnson, J. D., Oberkampf, W. L., and Storlie, C. B., "A sampling-based computational strategy for the representation of epistemic uncertainty in model predictions with evidence theory," Tech. Rep. SAND2006-5557, Sandia National Laboratories, 2006.

[24]Kleb, W. L. and Johnston, C. O., "Uncertainty analysis of air radiation for lunar return shock layers," *AIAA Atmospheric Flight Mechanics Conference and Exhibit*, Honolulu, HI, August 2008, AIAA Paper, 2008-6388.

[25]Diegert, K., Klenke, S., Novotny, G., Paulsen, R., Pilch, M., and Trucano, T., "Toward a More Rigorous Application of Margins and Uncertainties within the Nuclear Weapons Life Cycle A Sandia Perspective," Tech. Rep. SAND2007-6219, Sandia National Laboratories, 2007.

[26]Hassan, B., *Thermo-Chemical Nonequilibrium Effects on the Aerothermodynamics of Hypersonic Vehicles*, Ph.D. thesis, North Carolina State University, 1993.

[27]Peter A. Gnoffo, R. N. G. and Shinn, J. L., "Conservation Equations and Physical Models for Hypersonic Air Flows in Thermal and Chemical Nonequilibrium," Tech. rep., NASA, February 1989.

[28]Blottner, F.G., J. M. and Ellis, M., "Chemically Reacting Gas Viscous Flow Program for Multi-Component Gas Mixtures," Tech. Rep. Sandia Release No. SC-RR-70-754, Sandia National Laboratories, Albuquerque, NM, December 1971.

[29]Wilke, C. R., "A Viscosity Equation for Gas Mixtures," *Journal of Chemical Physics*, Vol. 18, No. 4, March 1950, pp. 517–519.

[30]Olynick, D. R., *A New LU-SGS Flow Solver for Calculating Reentry Flows*, Ph.D. thesis, North Carolina State University, 1992.

[31]Edwards, J. R., "A Low-Diffusion Flux-Splitting Scheme for Navier-Stokes Calculations," *Computers & Fluids*, Vol. 26, No. 6, 1997, pp. 635–659.

[32]Cheatwood, F. M. and Gnoffo, P. A., "User's Manual for the Langley Aerothermodynamic Upwind Relaxation Algorithm (LAURA)," Tech. rep., NASA, April 1996.

[33]NASA, *FUN3D: Fully Unstructured Navier-Stokes Manual*, May 2009, http://fun3d.larc.nasa.gov/index.html.

[34]Cheatwood, F. M. and Gnoffo, P. A., "User's Manual for the Langley Aerothermodynamic Upwind Relaxation Algorithm (LAURA)," NASA Technical Memorandum, 1996.

[35]Lockwood, B. A. and Mavriplis, D. J., "Parameter Sensitivity Analysis for Hypersonic Viscous Flow Using a Discrete Adjoint Approach," *48th AIAA Aerospace Sciences Meeting and Exhibit*, Orlando, FL, January 2010, AIAA Paper, 2010-447.

[36]Mavriplis, D., "Formulation and Multigrid Solution of the Discrete Adjoint for Optimization Problems on Unstructured Meshes," *43rd AIAA Aerospace Sciences Meeting*, Reno, NV, January 2005, AIAA Paper 2005-0319.

[37]Hascoët, L., "TAPENADE: a tool for Automatic Differentiation of programs," *Proceedings of 4^{th} European Congress on Computational Methods, ECCOMAS'2004, Jyvaskyla, Finland*, 2004.

[38]Wright, M. J., Bose, D., and Chen, Y.-K., "Probabilistic modeling of aerothermal and thermal protection material response uncertainties," *AIAA Journal*, Vol. 45(2), 2007, pp. 399425.

[39]Kreinovich, V. and Ferson, S., "A New Cauchy-Based Black-Box Technique for Uncertainty in Risk Analysis," *Reliability Engineering and Systems Safety*, 2002, pp. 267–279.

[40]Sherman, L. L., Taylor III, A. C., Green, L. L., and Newman, P. A., "First- and second-order aerodynamic sensitivity derivatives via automatic differentiation with incremental iterative methods," *Journal of Computational Physics*, Vol. 129, 1996, pp. 307 – 331.

[41]Putko, M. M., Newmann, P. A., Taylor III, A. C., and Green, L. L., "Approach for uncertainty propagation and robust design in CFD using sensitivity derivatives," *15th AIAA Computational Fluid Dynamics Conference*, Anaheim,CA, June 2001, AIAA Paper, 2001-2528.

[42]Rumpfkeil, M. P. and Mavriplis, D. J., "Efficient Hessian Calculations using Automatic Differentiation and the Adjoint Method with Applications," *AIAA Journal*, Vol. accepted, 2010.

[43]Chalot, F., Dinh, Q., Herbin, E., Martin, L., Ravachol, M., and Roge, G., "Estimation of the impact of geometrical uncertainties on aerodynamic coefficients using CFD," *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Schaumburg, IL, April 2008, AIAA Paper, 2068-2008.

[44]Rumpfkeil, M. P., Yamazaki, W., and Mavriplis, D. J., "Uncertainty Analysis Utilizing Gradient and Hessian Information," *Sixth International Conference on Computational Fluid Dynamics, ICCFD6*, St. Petersburg, Russia, July 2010.

[45]Pierce, N. A. and Giles, M. B., "Adjoint and defect error bounding and correction for functional estimates," *Journal of Computational Physics*, Vol. 200 No. 2, 2004, pp. 769–794.

[46]Giles, M. B., Pierce, N. A., and Sueli, E., "Progress in adjoint error correction for integral functionals," *Computing and Visualization in Science*, Vol. 6(2-3), 2004, pp. 113–121.

[47]Yager, R. R. and Liu, L., *Classic works of the Dempster-Shafer theory of belief functions. Studies in fuzziness and soft computing series, v. 219*, Berlin: Springer, 2008.

[48]Alina Alexeenko, A.B. Weaver, R. G. and Camberos, J., "Flowfield Uncertainty Analysis for Hypersonic CFD Simulations," *48th AIAA Aerospace Sciences Meeting and Exhibit*, Orlando, FL, January 2010, AIAA Paper, 2010-1180.

[49]Bil Kleb, R. A. T. and Johnston, C. O., "Blurring the Inputs: A Natural Language Approach to Sensitivity Analysis," *18th AIAA Computational Fluid Dynamics Conference*, Miami, FL, June 2007, AIAA Paper, 2007-4206.

[50]Lewis, E. and Miller, W., *Computaional Methods of Neutron Transport*, American Nuclear Society, 1993.

[51]"Variance," December 2010, http://en.wikipedia.org/wiki/Variance, Last modified 17 December 2010.

[52]"Kurtosis," December 2010, http://en.wikipedia.org/wiki/Kurtosis, Last modified 14 December 2010.